AMENDMENT TO THE CLAIMS

1. (currently amended) A method for determining a word entered using a reduced keypad, where each of one or more keys of the reduced keypad is mapped to a plurality of letters, the method comprising:

receiving key input corresponding to the entered word and~~,~~ ~~the key input having~~ at least one of a left context and a right context; ~~and~~

determining a list of possible words~~the word~~ corresponding to the key input for the entered word, wherein each listed word is in a vocabulary or a cache; and ~~by~~

using a ~~machine learning approach with a~~ language model to rank the listed words based on one or more of the at least one of the left context and the right context of the key input, wherein the language model is trained using words entered into the cache.

2. (previously presented) The method of claim 1, wherein the reduced keypad is a numeric keypad.

3. (currently amended) The method of claim 1, wherein the key input has at least the left context, and the word corresponding to the key input is determined ~~by using the machine learning approach~~ based in part on the left context of the key input.

4. (currently amended) The method of claim 1, wherein the key input has at least the right context, and the word corresponding to the key input is determined ~~by using the machine learning approach~~ based in part on the right context of the key input.

5. (currently amended) The method of claim 1, wherein the key input has both the left context and the right context, and the

word corresponding to the key input is determined ~~by using the machine learning approach~~ based both on the left context and the right context of the key input.

6. (currently amended) The method of claim 1, wherein using the language model comprises using a word~~an~~ n-gram model.

7. (original) The method of claim 6, wherein using the n-gram model comprises using a bigram model.

8. (currently amended) The method of claim 1, wherein using the language model~~machine learning approach~~ further comprises using a cache model ~~in addition to the language model~~.

9. (original) The method of claim 1, wherein the language model comprises a compressed language model.

10. (currently amended) The method of claim 1, wherein the key input has at least the left context, and wherein the word corresponding to the key input is determined ~~by using the machine learning approach~~ based in part on the left context of the key input, and wherein using the language model ~~machine learning approach~~ comprises using a bigram model as the language model, comprising the steps:

for each listed word ~~in a vocabulary that is consistent with the key input,~~ determining a probability of the word given the left context, and adding the word and the probability of the word to an array of word-probability pairs; and~~,~~

sorting the array of word-probability pairs in decreasing order of probability.

11. (currently amended) The method of claim 10, wherein using the language model comprises ~~machine learning approach comprises~~ using a cache model in addition to using the bigram model, such that the probability of the word is determined given the left context and the words entered in the~~a~~ cache.

12. (original) The method of claim 10, wherein using the bigram model further comprises:

for each word in the vocabulary that is consistent with the key input as an initial part of the word, determining a probability of the word given the left context, and, upon determining that the probability is greater than a greatest probability so far determined, setting the greatest probability to the probability and a greatest probability word associated with the greatest probability to the word;

upon determining that the greatest probability is at least a number of times greater than a word of a first word-probability pair of the array of word-probability and the greatest probability as a new first word-probability pair before the first word-probability pair within the array.

13. (currently amended) The method of claim 12, wherein using the language model ~~machine learning approach~~ comprises using a cache model in addition to using the bigram model, such that the probability of the word is determined given the left context and words entered in the~~a~~ cache.

14. (currently amended) The method of claim 1, wherein the key input has both the left context and the right context and has a plurality of number sequences where each sequence corresponds to a word, the plurality of words corresponding to the key input

determined by using the ~~machine learning approach~~ language model based in part on both ~~on~~ the left context and the right context of the key input.

15. (original) The method of claim 1, wherein the language model comprises a compressed language model, the compressed language model compressed by performing a method comprising:

smoothing an uncompressed language model; and,

pruning the uncompressed language model to yield the compressed language model.

16. (original) The method of claim 15, wherein pruning the uncompressed language model comprises using one of: count-cutoffs approach, a Rosenfeld pruning approach, and a Stolcke pruning approach.

17. (original) The method of claim 15, wherein pruning the uncompressed language model comprises determining a normalization factor for each word in the uncompressed model only prior to pruning.

18. (original) The method of claim 15, wherein pruning the uncompressed language model accounts for ambiguous words in the uncompressed model.

19. (original) The method of claim 15, wherein pruning the uncompressed language model accounts for an effect of the pruning on key input accuracy.

20. (original) The method of claim 1, wherein the method is performed by execution of a computer program by a processor from a computer-readable medium.

21. (currently amended) A computer-readable medium having instructions stored thereon for execution by a processor to perform a method for determining a word entered using a reduced keypad, where each of one or more input keys of the reduced keypad is mapped to a plurality of letters, the method comprising:

receiving key input corresponding to the word and~~, the key input having~~ a left context;

for each word in a vocabulary that is consistent with the key input, determining an n-gram~~a~~ probability of the word given the left context, and adding the word and the n-gram~~probability~~ of the word to an array of word-probability pairs, wherein the n-gram probabilities are stored in a language model trained at least in part on words entered into a cache;

determining the word corresponding to the key input as a word of a word-probability pair within the array of word-probability pairs having a greatest probability.

22. (original) The medium of claim 21, wherein the reduced keypad is a numeric keypad.

23. (original) The medium of claim 21, wherein determining the word corresponding to the key input comprises:

sorting the array of word-probability pairs in decreasing order of probability; and

determining the word corresponding to the key input as a word of a first word-probability pair within the array of word-probability pairs.

24. (currently amended) The medium of claim 21, the method further initially comprising, for each word in the~~a~~ cache that is consistent with the key input, determining a probability of the

word given the left context, and adding the word and the probability of the word to an array of word-probability pairs.

25. (original) The medium of claim 21 the method further comprising prior to determining the word corresponding to the key input:

for each word in the vocabulary that is consistent with the key input as an initial part of the word, determining a probability of the word given the left context, and, upon determining that the probability is greater than a greatest probability so far determined, setting the greatest probability to the probability and a greatest probability word associated with the greatest probability to the word;

upon determining that the greatest probability is significantly more likely than a word of a first word-probability pair of the array of word probability-pairs, adding the greatest probability word associated with the greatest probability and the greatest probability as a new first word-probability pair to the array.

26. (currently amended) The medium of claim 25 the method further initially comprising prior to determining the word corresponding to the key input, for each word in thea~~—~~ cache that is consistent with the key input as an initial part of the word, determining a probability of the word given the left context, and, upon determining that the probability is greater than the greatest probability so far determined, setting the greatest probability to the probability and a greatest probability word associated with the greatest probability to the word.

27. (currently amended) A method for determining a word entered using a reduced keypad, where each of one or more keys of the reduced keypad is mapped to a plurality of letters, the method comprising:

receiving key input corresponding to the word~~, the key input having~~ and at least one of a left context and a right context;

determining the word corresponding to the key input by using a compressed language model based on one or more of the at least one of the left context and the right context of the key input, wherein the language model is trained in part using words entered in a cache, and wherein the ~~compressed~~ language model is compressed by performing the steps of~~a method comprising~~:

smoothing the~~an~~ uncompressed language model; and,

pruning the uncompressed language model to yield the compressed language model.

28. (original) The method of claim 27, wherein the reduced keypad is a numeric keypad.

29. (original) The method of claim 27, wherein pruning the uncompressed language model comprises using one of: a count-cutoffs approach, a Rosenfeld pruning approach, and a Stolcke pruning approach.

30. (original) The method of claim 27, wherein pruning the uncompressed language model comprises determining a normalization factor for each word in the uncompressed model only prior to pruning.

31. (original) The method of claim 27, wherein pruning the uncompressed language model accounts for ambiguous words in the uncompressed model.

32. (original) The method of claim 27, wherein pruning the uncompressed language model accounts for an effect of the pruning on key input accuracy.

33. (original) The method of claim 27, wherein the method is performed by execution of a computer program by a processor from a computer-readable medium.

34. (currently amended) An apparatus comprising:

a plurality of keys, each of one or more of the keys mapped to a plurality of letters, the plurality of keys used to enter key input corresponding to a word and~~, the key input having~~ at least one of a left context and a right context; and,

a word-determining logic designed to construct a list of possible words corresponding to the entered word and ranking the listed words to determine the word corresponding to the key input by using ~~a machine learning approach with~~ a language model based on one or more of the at least one of the left context and the right context of the key input, wherein the language model is further based on words entered into a cache.

35. (original) The apparatus of claim 34, further comprising a display on which the at least one of the left context and the right context, and the word corresponding to the key input, are displayed.

36. (original) The apparatus of claim 34, wherein the apparatus is a telephone.

37. (original) The apparatus of claim 36, wherein the apparatus is a mobile telephone.

38. (original) The apparatus of claim 36, wherein the apparatus is one of: a cellular telephone, a corded telephone, a cordless telephone, a digital telephone, and a radio telephone.

39. (original) The apparatus of claim 34, wherein the apparatus is one of: a pager, a desktop computer, a laptop computer, a handheld device, a personal-digital assistance (PDA) device, and a remote control device.

40. (original) The apparatus of claim 34, wherein the word-determining logic comprises a computer program stored on a computer-readable medium for execution by a processor.

41. (currently amended) The apparatus of claim 34, wherein the key input has at least the left context, and the word corresponding to the key input is determined by the word-determining logic by training the language model~~using the machine learning approach~~ based in part on the left context of the key input.

42. (currently amended) The apparatus of claim 34, wherein the key input has at least the right context, and the word corresponding to the key input is determined by the word-determining logic by training the language model~~using the machine learning approach~~ based in part on the right context of the key input.

43. (currently amended) The apparatus of claim 34, wherein the key input has both the left context and the right context, and the word corresponding to the key input is determined by the word-determining logic by training the language model ~~using the machine learning approach~~ based in part on both ~~on~~ the left context and the right context of the key input.

44. (original) The apparatus of claim 34, wherein the word-determining logic uses a cache model.